



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/659,983
Source: Oise
Date Processed by STIC: 9-22-03

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>), EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 04/24/2003

Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u>	<u>SUGGESTED CORRECTION</u>	<u>SERIAL NUMBER:</u> <u>10/659,983</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 <input type="checkbox"/> Wrapped Nucleic Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 <input type="checkbox"/> Misaligned Amino Numbering	The numbering under each 5 th amino acid is misaligned. Do not use tab codes between numbers; use space characters , instead.	
4 <input type="checkbox"/> Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 <input type="checkbox"/> Variable Length	Sequence(s) <input type="checkbox"/> contain n's or Xaa's representing more than one residue. Per Sequence Rules , each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) <input type="checkbox"/> . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) <input type="checkbox"/> missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) <input type="checkbox"/> missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <input type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 <input checked="" type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 <input type="checkbox"/> Use of <220>	Sequence(s) <input type="checkbox"/> missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 07/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 <input type="checkbox"/> Misuse of n/Xaa	"n" can only represent a single <u>nucleotide</u> ; "Xaa" can only represent a single <u>amino acid</u>	



OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/659,983

DATE: 09/22/2003
TIME: 14:19:07

Input Set : A:\81289-284779.ST25.txt
Output Set: N:\CRF4\09222003\J659983.raw

```

3 <110> APPLICANT: Hovanec, Timothy A
5 <120> TITLE OF INVENTION: Ammonia-Oxidizing Bacteria
7 <130> FILE REFERENCE: 81289-284779
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/659,983
C--> 9 <141> CURRENT FILING DATE: 2003-09-10
9 <150> PRIOR APPLICATION NUMBER: US 09/573,684
10 <151> PRIOR FILING DATE: 2000-05-19
12 <150> PRIOR APPLICATION NUMBER: US 60/386,217
13 <151> PRIOR FILING DATE: 2002-09-19
15 <150> PRIOR APPLICATION NUMBER: US 60/386,218
16 <151> PRIOR FILING DATE: 2002-09-19
18 <150> PRIOR APPLICATION NUMBER: US 60/386,219
19 <151> PRIOR FILING DATE: 2002-09-19
21 <160> NUMBER OF SEQ ID NOS: 23
23 <170> SOFTWARE: PatentIn version 3.2
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 1457
27 <212> TYPE: DNA
28 <213> ORGANISM: AOB Type A R7clone140 16S rDNA (SEQ ID NO:1)
30 <400> SEQUENCE: 1
31 attgaacgct ggcggcatgc tttacacatg caagtcaaac ggcagcacgg atgttgcatt 60
33 ctgggtggcga gtggccgacg ggtggataat gcatcgaaac gtatccagaa gaggggggta 120
35 acgcacatcgaa agatgtgcta ataccgcata tactctaagg aggaaagcag gggatcgaaa 180
37 gacccttgcgc ttttggagcg gccgatgtct gattagctag ttgggtgggt aaaggcctac 240
39 caaggcgcacg atcagtagtt ggtctgagag gacgaccgcg cacactggga ctgagacacg 300
41 gcccagactc ctacgggagg cagcagtgg gaattttgga caatgggcgc aagcctgatc 360
43 cagaatgcgc gcgtgagtga agaaggcctt cgggtttaa agctcttca gtcgagaaga 420
45 aaagggttacg gtaaaataatc gtgactcatg acggtatcga cagaagaagc accggctaac 480
47 tacgtgccag cagccgcgtt aatacgtagg gtgcaagcgt taatcgaaat tactggcgt 540
49 aaagggttgcg caggccgctt tgtaagttag atgtgaaatc cccgggctta acctggaaat 600
51 tgcgtttgaa actacaaggc tagagtgtgg cagagggagg tggatttcca tgttagcag 660
53 tggaaatgcgt agagatatgg aagaacatcg atggcaagg cagcctctg ggttaacact 720
55 gacgctcatg cacgaaagcg tggggagcaa acaggattag atacccttgtt agtccacgcc 780
57 ctaaaacgtatg tcaacttagtt gttgggcctt attaggctt gtaacgaagc taacgcgtga 840
59 agttgaccgc ctggggagta cggcgcgaag attaaaactc aaaggaattt acggggaccc 900
61 gcacaaggcg tggattatgt ggattaattc gatgcaacgc gaaaaacacctt acctaccctt 960
63 gacatgtacg gaattttcta gagatagatt agtgcgttcgg gaacgcatac acaggtgctg 1020
65 catggctgtc gtcagctcgt gtcgtgagat gttgggttaa gtcccgcaac gagcgcaccc 1080
67 cttgtcatta attgccccatca tttgggttggg cacttaatg agactgcgg tgacaaaccg 1140
69 gaggaagggtt gggatgacgt caagtccctca tggcccttat gggtagggct tcacacgtaa 1200
71 tacaatggcg cgtacagagg gttgccaacc cgccgaggggg agctaatttc agaaagcgcg 1260
73 tcgttagtccg gatcgagatc tgcaactcga ctccgtgaag tcggaatcgc tagtaatcgc 1320
75 ggatcagcat gtcgcgttga atacgttccc gggcttgta cacaccgccc gtcacaccat 1380

```

**Does Not Comply
Corrected Diskette Need d**

See item 10
on error summary
sheet.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/659,983

DATE: 09/22/2003
TIME: 14:19:07

Input Set : A:\81289-284779.ST25.txt
Output Set: N:\CRF4\09222003\J659983.raw

77	gggatgggt ttcaccagaa gcaggtagtc taaccgtaa gagggcgctt gccacggta	1440
79	gattcatgac tggggtg	1457
82	<210> SEQ ID NO: 2	
83	<211> LENGTH: 1457	
84	<212> TYPE: DNA	
85	<213> ORGANISM: AOB Type A1 R7clone187 16S rDNA (SEQ ID NO:2)	
87	<400> SEQUENCE: 2	
88	attgaacgct ggcggcatgc ttacacatg caagtcgaac ggcagcacgg atgcttgcat	60
90	ctgggtggcga gtggcggacg ggtgagtaat gcatcgaaac gtatccagaa gaggggggta	120
92	acgcacatcgaa agatgtgcta ataccgcata tactctaagg aggaaagcag gggatcgaaa	180
94	gaccttgcgc ttttggagcg gcccgtgtct gattagctag ttgggtgggt aaaggcctac	240
96	caaggcgacg atcagttagt ggctctgagag gacgaccacg cacactggga ctgagacacg	300
98	gcccgacactc ctacgggagg cagcagtggg gaattttgga caatggcgc aagcctgatc	360
100	cagcaatgcc gcgtgagtgaa agaaggcctt cggttgtaa agctcttca gtcgagaaga	420
102	aaagggttacg gtaaataatc gtgaccatcg acggtatcgaa cagaagaagc accggtaac	480
104	tacgtgccag cagccgcgtt aatacgttagg gtgcagcgt taatcggat tactggcgt	540
106	aaagggtgcg caggcggcct tctaagtcaatg atgtgaaatc cccgggctt acctggaaat	600
108	tgcgtttgaa actacaatcg tagagtgtgg cagaggagg tggaaattcca tggtagcag	660
110	tgaaatcgat agagatatgg aagaacatcg atggcgaagg cagcctcctg gtttaacact	720
112	gacgctcatg cacgaaagcg tggggagcaa acaggattag ataccctgtt agtccacgccc	780
114	ctaaacatcg tcaacttagt gttgggcctt attagctt gtaacgaagc taacgcgtga	840
116	attgaccgcg ctggggagta cggtcgcaag attaaaaactc aaaggaattt acggggaccc	900
118	gcacaagcgg tggattatgt ggattaattt gatcAACGC gaaaaacctt acctaccctt	960
120	gacatgttagc gaattttcta gagatagatt agtgcctcg gAACGCTAAC acagggtctg	1020
122	catggctgtc gtcagctcgat gtcgtgagat gttgggttaa gtcccgcaac gagcgcaccc	1080
124	cttgcattt attgcattca ttgggtggg cactttatg agactgcgg tgacaaaccg	1140
126	gaggaagggtt gggatgacgt caagtctca tggcccttat gggttagggct tcacacgtaa	1200
128	tacaatggcg cgtacagagg ttgcacacc cgcgggggg agctaatttc agaaagcgcg	1260
130	tctgtgtccg gatcgaggatc tgcaactcgat ctccgtgaag tggaaatcgc tagtaatcgc	1320
132	ggatcagcat gtcgcgtgtaa atacgttccc gggcttggta cacaccgcac gtcacaccat	1380
134	gggatgggtt ttcaccagaa gcaggtagtc taaccgtaa gagggcgctt gccacggta	1440
136	gattcatgac tggggtg	1457
139	<210> SEQ ID NO: 3	
140	<211> LENGTH: 1458	
141	<212> TYPE: DNA	
142	<213> ORGANISM: AOB Type B R3clone5 16S rDNA (SEQ ID NO:3)	Same
144	<400> SEQUENCE: 3	
145	attgaacgct ggcggcatgc ttacacatg caagtcgaac ggcagcacgg gggcaaccct	60
147	gtggcgagt ggcgaacggg ttagtaatac atcgaaacgt atcttcgagg gggggataac	120
149	gcaccgaaat gttgtcaat accgcataat ctccacggag aaaagcagg gatcgaaga	180
151	ccttgcgttc ttggagcggc cgtatgtctga ttagcttagt ggtgaggtaa tggcttacca	240
153	aggcgacgt cagtagctgg tctgagagga cgaccagcca cactggact gagacacggc	300
155	ccagactcct acgggaggca gcaatggggg atttggaca atggggaaa ccctgtatcca	360
157	gcatgcccgc gtgagtgaaat aaggccttgg gttgtaaat ctcttcagc cggaacgaaa	420
159	cgttcacggc taataccgt gactactgac ggtaccggaa gaagaagcac cggtacta	480
161	cgtgccagca gcccggtaa tacgttaggt gcaagcgtt atcggaaatta ctggcgtaa	540
163	acgtgcgcgca ggcgggtttt taagtcaat gtaacggccc cgggcttaac ctgggaactg	600
165	cgttggaaac tacaaggcta ggtgtggca gaggggggtt gatattccacg tggtagcagt	660
167	aaatqcgtaa aqatqggqa qaacaccgat ggcgaaggca gccccctggg ttaacaccga	720

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/659,983

DATE: 09/22/2003

TIME: 14:19:07

Input Set : A:\81289-284779.ST25.txt

Output Set: N:\CRF4\09222003\J659983.raw

169	cgctcaggca	cgaaaagctg	gggagcaaaac	aggatttagat	accctggtag	tccacgccc	780
171	aaacgatgtc	aactagtgt	cgggtcttaa	cggaacttgc	aacgcagcta	acgcgtgaag	840
173	ttggccgcct	ggggagtacg	gtcgcaagat	taaaactcaa	aggaatttgc	ggggacccgc	900
175	acaaggcggt	gattatgtgg	attaattcga	tgcaacgcga	aaaaccttac	ctacccttga	960
177	catgtaccga	agccccggca	gagggtgggt	tgcccgaaag	ggagcggtaa	cacaggtgt	1020
179	gcatggctgt	cgtcagctcg	tgtcggtgaga	tgttgggtt	agtcccgaa	cgagcgcaac	1080
181	ccttgtcatt	aattgccatc	attcagttgg	gcactttaat	gaaaactgcc	gtgacaaacc	1140
183	ggaggaagt	ggggatgacg	tcaagtcctc	atggccctta	tgggtagggc	ttcacacgt	1200
185	atacaatggc	gcgtacagag	gttgcacaa	ccgcgagggg	gagctaata	cagaaggcgc	1260
187	gtcgtagtcc	ggatcggt	ctgcaactcg	actccgtgaa	gtcggaaatcg	ctagtaatcg	1320
189	cgatcgac	tgtcggtg	aatacgttcc	cgggtcttgc	acacaccgc	cgtcacaccca	1380
191	tgggagtgg	tttaccaga	agcaggtgt	ctaaccgaa	ggagggcgct	tgccacgggt	1440
193	agattcatga	ctgggtg					1458
196	<210>	SEQ ID NO:	4				
197	<211>	LENGTH:	1460				
198	<212>	TYPE:	DNA				
199	<213>	ORGANISM:	AOB Type C R5clone47 16S rDNA (SEQ ID NO:4)	Same			
201	<400>	SEQUENCE:	4				
202	attgaacgct	ggcgccatgc	tttacacatg	caagtcgaac	ggcagcgccc	gttcggc	60
204	gccggcgagt	ggcgaacggg	ttagtaatac	atcggaacgt	gtccttaagt	gggataaac	120
206	gcatcgaaag	atgtgcta	accgcata	tctgaggaga	aaagcagggg	atcgcaagac	180
208	cttgcgtt	aaaggacggcc	gatgtctgt	tagctagtt	tgtgggtaaa	ggcttaccaa	240
210	ggcaacgatc	agtatgtgg	ctgagaggac	gaccaaccac	actgggact	agacacggcc	300
212	cagactccta	cgggaggcag	cagtggggaa	ttttggacaa	tggcgaaag	cctgatccag	360
214	ccatgcccgc	ttagtgaaga	aggccttcgg	ttttagagc	tcttttagtc	agaaagaaag	420
216	aatcatgt	aataattatg	atttatgac	gtactgac	aaaaagcacc	ggctaactac	480
218	gtgccagcag	ccgcgtt	acgttaggt	cgagcgtt	tccgaaattac	tggcgtaaa	540
220	gggtgcgcag	gcgggtt	taagtca	tgaaagcccc	ggcttaacc	tggaaattgc	600
222	tttggaaact	acaaggctag	agtgcagcag	aggggagtgg	aattccatgt	gtacgtgt	660
224	aatgcgtaga	gatgtggaa	aacaccgtat	gcgaaggcag	ctccctgggt	tgacactgac	720
226	gctcatgcac	gaaagcgtgg	ggagcaaa	ggattagata	ccctggtagt	ccacgccc	780
228	aacgatgtca	actgttgc	ggatctaatt	aaggatttg	taacgtact	aacgcgtgaa	840
230	gttgaccggcc	tggggagtac	ggtcgcaaga	ttaaaactca	aaggaattga	cggggacccg	900
232	cacaaggcgt	ggattatgt	gattaattcg	atgcaacgc	aaaaaccta	cctaccctt	960
234	acatgcgtt	aatctatgt	agacataaga	gtgccc	gggagcca	acacagggtc	1020
236	tgcgttgc	tgcgttgc	gtgtcggt	atgttgggt	aagtccgc	acgagcgca	1080
238	cccttgcac	taattgtat	cattctaa	gagactt	gtgagact	cggtgacaaa	1140
240	ccggaggaa	gtgggatga	cgtcaagtcc	tatggcc	tatgggt	gttcacacg	1200
242	taatacaatg	gcgtgtac	agggttgc	accgcgagg	gggagcc	atctcagaa	1260
244	acgtcgtagt	ccggatcg	gtctgcact	cgactccgt	aagtccgaa	cgctagtaat	1320
246	cgccggatcag	catggccg	tgaatacgtt	ccgggtt	gtacacacc	cccgacac	1380
248	catgggagt	gtttcacca	gaagcaggta	gtttaaccgt	aaggaggac	cttgcac	1440
250	tgggggtcat	gactgggt					1460
253	<210>	SEQ ID NO:	5				
254	<211>	LENGTH:	18				
255	<212>	TYPE:	DNA				
256	<213>	ORGANISM:	Oligonucleotide Probe (SEQ ID NO:5)	See Item 10 On error summary report			
258	<400>	SEQUENCE:	5				
259	ccccctt	ctggata					

18

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/659,983

DATE: 09/22/2003
TIME: 14:19:07

Input Set : A:\81289-284779.ST25.txt
Output Set: N:\CRF4\09222003\J659983.raw

262 <210> SEQ ID NO: 6	
263 <211> LENGTH: 18	
264 <212> TYPE: DNA	
265 <213> ORGANISM: PCR primer (SEQ ID NO:6)	
267 <400> SEQUENCE: 6	18
268 cggaacgtat ccagaaga	
271 <210> SEQ ID NO: 7	
272 <211> LENGTH: 18	
273 <212> TYPE: DNA	
274 <213> ORGANISM: PCR primer (SEQ ID NO:7)	
276 <400> SEQUENCE: 7	18
277 atctctagaa aattcgct	
280 <210> SEQ ID NO: 8	
281 <211> LENGTH: 19	
282 <212> TYPE: DNA	
283 <213> ORGANISM: Oligonucleotide probe (SEQ ID NO:8)	
285 <400> SEQUENCE: 8	19
286 tcccccactc gaagatacg	
289 <210> SEQ ID NO: 9	
290 <211> LENGTH: 17	
291 <212> TYPE: DNA	
292 <213> ORGANISM: PCR primer (SEQ ID NO:9)	
294 <400> SEQUENCE: 9	17
295 atcggAACGT atcttcg	
298 <210> SEQ ID NO: 10	
299 <211> LENGTH: 16	
300 <212> TYPE: DNA	
301 <213> ORGANISM: PCR primer (SEQ ID NO:10)	
303 <400> SEQUENCE: 10	16
304 ccacctctcr gcgggc	
307 <210> SEQ ID NO: 11	
308 <211> LENGTH: 19	
309 <212> TYPE: DNA	
310 <213> ORGANISM: PCR primer (SEQ ID NO:11)	
312 <400> SEQUENCE: 11	19
313 tcagaaagaa agaatcatg	
316 <210> SEQ ID NO: 12	
317 <211> LENGTH: 19	
318 <212> TYPE: DNA	
319 <213> ORGANISM: PCR primer (SEQ ID NO:12)	Same error
321 <400> SEQUENCE: 12	19
322 gtctccayta gattccaag	
325 <210> SEQ ID NO: 13	
326 <211> LENGTH: 17	
327 <212> TYPE: DNA	
328 <213> ORGANISM: PCR primer (SEQ ID NO:13)	
330 <400> SEQUENCE: 13	17
331 gtttgcaccc ggctcag	
334 <210> SEQ ID NO: 14	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/659,983

DATE: 09/22/2003
TIME: 14:19:07

Input Set : A:\81289-284779.ST25.txt
Output Set: N:\CRF4\09222003\J659983.raw

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/659,983

DATE: 09/22/2003
TIME: 14:19:08

Input Set : A:\81289-284779.ST25.txt
Output Set: N:\CRF4\09222003\J659983.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date